

---

## **NLSY97 Appendix 4: Geographic Variable Creation**

---



Several variables in the main data set provide information about the respondent's area of residence. These variables are intended to permit researchers to identify key characteristics of the area without needing access to the geocode CD-ROM. Geographic variables were created using a software program called Matchmaker for Windows, V2.5; therefore, no programming code is provided for these variables. Instead, this document offers a brief general description of the methods used to generate these variables. For more information about the process of classifying a respondent's metropolitan area or about the geographic variables in general, refer to the introduction to the *Geocode Codebook Supplement* or contact NLS User Services.

## CENSUS REGION OF RESIDENCE AT SURVEY DATE AND AT AGE 12

Variables Created: CV\_CENSUS\_REGION  
CV\_CENSUS\_REGION\_AGE\_12

These variables classify respondents as residing in one of four regions defined by the U.S. Bureau of the Census. These regions are as follows:

Census Division	States
Northeast	Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont
North Central	Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Dakota, Ohio, South Dakota, and Wisconsin
South	Alabama, Arkansas, Delaware, District of Columbia, Florida, Georgia, Kentucky, Louisiana, Maryland, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia
West	Alaska, Arizona, California, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, Wyoming

## MSA STATUS AT SURVEY DATE AND AT AGE 12

Variables Created: CV\_MSA  
CV\_MSA\_AGE\_12

This variable provides users with information about whether the respondent lived in the central city of the MSA, in another part of the MSA, or outside of an MSA. As defined by the Census Bureau, a central city is the major city lying within a Metropolitan Statistical Area (MSA). Initially, two variables were created using the TIGER/Line files (a database developed by the Census Bureau) to classify respondent residences: one identifying the respondent's MSA status and one identifying the respondent's central city status. The variables were then combined to produce a single MSA/central city variable. For this dataset, respondents are coded as follows:

- 1 not in MSA
- 2 in MSA, not central city
- 3 in MSA, central city
- 4 in MSA, not known
- 5 not in country

## **RURAL VS. URBAN**

**Variables Created:** CV\_URBAN\_RURAL  
CV\_URBAN\_RURAL\_AGE\_12

As defined by the geocode software, urban places are “closely settled, named, communities that generally contain a mixture of residential, commercial, and retail areas, and have a population greater than 2,500.” These criteria are based on those used by the Census Bureau to identify urban areas. All other places are considered rural. Based on the TIGER/Line files, respondents residing in urban areas are coded 1 and those residing in rural areas are coded 0. Census Bureau information on urban and rural places can be retrieved from the following internet site:

<http://www.census.gov/geo/www/tiger/index.html>